

# DEVICE AND METHOD FOR PREPARING CHINESE DOCUMENT

Publication number: JP2001109740

Publication date: 2001-04-20

Inventor: ISHIZUKA YASUSHI; KAMINARI KAITO

Applicant: TOKYO SHIBAURA ELECTRIC CO

Classification:

- international: G06F3/16; G06F17/21; G06F17/22; G06F3/16;  
G06F17/21; G06F17/22; (IPC1-7): G06F17/21;  
G06F3/16; G06F17/22

- European:

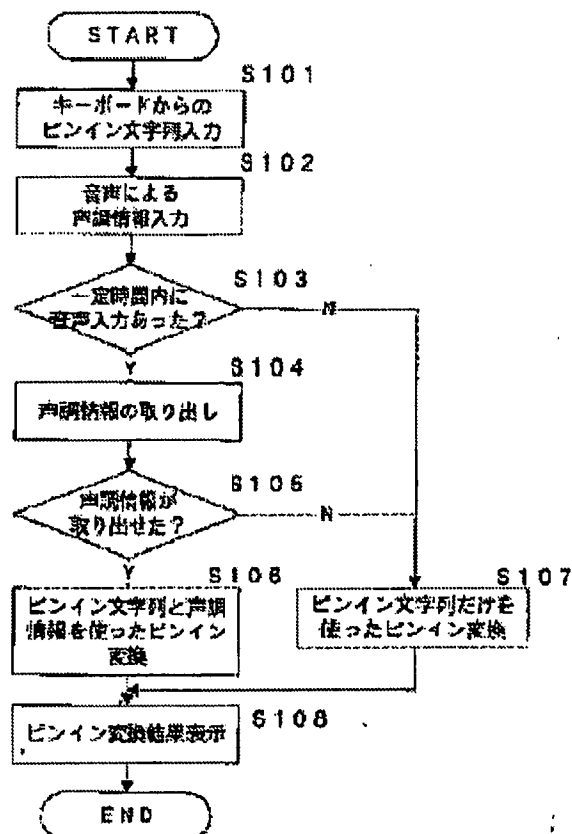
Application number: JP19990285050 19991006

Priority number(s): JP19990285050 19991006

Report a data error here

## Abstract of JP2001109740

**PROBLEM TO BE SOLVED:** To improve input efficiency by inputting a pinyin character string from a keyboard as in the conventional way and also inputting voice information by using voice. **SOLUTION:** A pinyin character string inputting means inputs the pinyin character string and also, a voice information inputting means inputs voice information of Kanji (Chinese character). A voice information analyzing means analyzes the inputted voice information and acquires tone information, and a pinyin converting means performs pinyin conversion on the basis of the tone information acquired by the voice information analyzing means and the pinyin character string inputted from the pinyin character string inputting means and acquires a conversion candidate.



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号  
特開2001-109740  
(P2001-109740A)

(43)公開日 平成13年4月20日(2001.4.20)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	デマコト* (参考)
G 0 6 F 17/21		G 0 6 F 3/16	3 2 0 H 5 B 0 0 9
3/16	3 2 0	15/20	5 9 2 J
17/22			5 0 3

審査請求 未請求 請求項の数6 O L (全 12 頁)

(21)出願番号 特願平11-285050

(22)出願日 平成11年10月6日(1999.10.6)

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 石塚 靖

東京都青梅市末広町2丁目9番地 株式会  
社東芝青梅工場内

(72)発明者 雷海濤

東京都青梅市末広町2丁目9番地 株式会  
社東芝青梅工場内

(74)代理人 100083161

弁理士 外川 英明

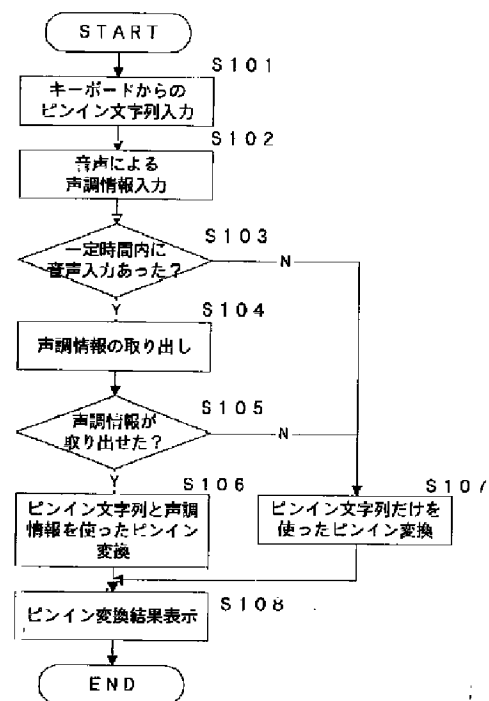
Fターム(参考) 5B009 KB01 ME05 ME21 MF02 MH01  
MH03 VB13

(54)【発明の名称】 中国語文書作成装置及び中国語文書作成方法

(57)【要約】

【課題】 ピンイン文字列は従来通りキーボードから入力すると共に、声調情報については音声を使って入力することにより入力効率の向上を可能とする。

【解決手段】 ピンイン文字列入力手段によってピンイン文字列を入力すると共に、音声情報入力手段によって漢字の音声情報を入力する。音声情報解析手段が、入力された音声情報を解析し声調情報を取得し、ピンイン変換手段が、音声情報解析手段によって取得された声調情報と、ピンイン文字列入力手段から入力されたピンイン文字列とに基づいてピンイン変換を行ない、変換候補を取得する。



## 【特許請求の範囲】

【請求項1】 ピンイン文字列を入力するピンイン文字列入力手段と、

ピンイン文字列と声調情報との組合せ及び当該組合せに対応する漢字をそれぞれ対応付けて格納するピンイン変換用辞書と、

前記入力手段から入力されるピンイン文字列に対応する漢字の音声情報を入力する音声情報入力手段と、

前記音声情報入力手段から入力された音声情報を解析し、前記ピンイン文字列入力手段から入力されたピンイン文字列に対応する声調情報を取得する音声情報解析手段と、

前記ピンイン文字列入力手段から入力された前記ピンイン文字列と前記音声情報解析手段によって取得された前記声調情報に基づいて前記ピンイン変換用辞書を参照し対応する漢字を取得するピンイン変換手段と、

前記ピンイン変換手段によって取得された漢字を変換候補として表示する表示手段とを具備することを特徴とする中国語文書作成装置。

【請求項2】 ピンイン文字列を入力するピンイン文字列入力手段と、

ピンイン文字列と声調情報との組合せ及び当該組合せに対応する漢字をそれぞれ対応付けて格納するピンイン変換用辞書と、

前記ピンイン文字列入力手段から入力された前記ピンイン文字列に基づいて前記ピンイン変換用辞書を参照し対応する漢字及び声調情報を取得するピンイン変換手段と、

前記ピンイン変換手段によって取得された漢字を変換候補として表示する表示手段と所望する漢字の音声情報を入力する音声情報入力手段と、

前記音声情報入力手段から入力された音声情報を解析し、前記ピンイン文字列入力手段から入力されたピンイン文字列に対応する声調情報を取得する音声情報解析手段と、

前記ピンイン変換手段によって取得された変換候補のうち前記音声情報解析手段によって取得された声調情報を持つ漢字を変換候補として優先的に表示するよう前記表示手段による変換候補の表示を制御する制御手段とを具備することを特徴とする中国語文書作成装置。

【請求項3】 ピンイン文字列を入力するピンイン文字列入力手段と、

ピンイン文字列と声調情報との組合せ及び当該組合せに対応する漢字をそれぞれ対応付けて格納するピンイン変換用辞書と、

前記入力手段から入力されるピンイン文字列に対応する漢字の音声情報を入力する音声情報入力手段と、

前記音声情報入力手段から入力された音声情報を解析し、前記ピンイン文字列入力手段から入力されたピンイン文字列に対応する声調情報を取得する音声情報解析手

段と、

前記ピンイン文字列入力手段から入力された前記ピンイン文字列と前記音声情報解析手段によって取得された声調情報に基づいて前記ピンイン変換用辞書を参照し対応する漢字を取得するピンイン変換手段と、

新たに登録されるユーザ単語に関するピンイン文字列と声調情報との組合せ及び当該組合せに対応する漢字をそれぞれ対応付けて格納するユーザ単語辞書と、

前記ユーザ単語辞書から前記ピンイン変換手段により取得された漢字と同一の漢字を持つユーザ単語を検索し、当該検索したユーザ単語の声調情報として前記音声情報解析手段によって取得された声調情報を登録するユーザ単語登録手段とを具備することを特徴とする中国語文書作成装置。

【請求項4】 ピンイン文字列を入力するピンイン文字列入力手段と、漢字の音声情報を入力する音声情報入力手段と、ピンイン文字列に対応する漢字を変換候補として表示する表示手段とを有する中国語文書作成装置における中国語文書作成方法において、

前記ピンイン文字列入力手段から入力されるピンイン文字列を受け付け、

前記音声情報入力手段から入力される音声情報を解析し、前記入力されたピンイン文字列に対応する声調情報を取得し、

ピンイン文字列と声調情報との組合せ及び当該組合せに対応する漢字をそれぞれ対応付けて格納するピンイン変換用辞書を参照し、前記受け付けたピンイン文字列と前記取得した声調情報に基づいて、前記受け付けたピンイン文字列に対応する漢字を取得し、

取得した漢字を変換候補として前記表示手段に表示させるよう制御することを特徴とする中国語文書作成装置。

【請求項5】 ピンイン文字列を入力するピンイン文字列入力手段と、漢字の音声情報を入力する音声情報入力手段と、ピンイン文字列に対応する漢字を変換候補として表示する表示手段とを有する中国語文書作成装置における中国語文書作成方法において、

前記ピンイン文字列入力手段から入力されるピンイン文字列を受け付け、

前記受け付けたピンイン文字列に基づいて、ピンイン文字列と声調情報との組合せ及び当該組合せに対応する漢字をそれぞれ対応付けて格納するピンイン変換用辞書を参照し、前記受け付けたピンイン文字列に対応する漢字及び声調情報を取得し、

取得した漢字を変換候補として前記表示手段に表示させるよう制御し、

変換候補の表示後に前記音声情報入力手段から入力される音声情報を解析し、前記入力されたピンイン文字列に対応する声調情報を取得し、

ピンイン変換により取得した変換候補のうち取得した声調情報と同一の声調情報を持つ漢字を変換候補として優

先に表示するよう前記表示手段による変換候補の表示を制御することを特徴とする中国語文書作成方法。

【請求項6】 ピンイン文字列を入力するピンイン文字列入力手段と、漢字の音声情報を入力する音声情報入力手段と、ピンイン文字列に対応する漢字を変換候補として表示する表示手段と、ユーザ単語を登録可能なユーザ単語辞書とを有する中国語文書作成装置における中国語文書作成方法において、

前記ピンイン文字列入力手段から入力されるピンイン文字列を受け付け、

前記音声情報入力手段から入力される音声情報を解析し、前記入力されたピンイン文字列に対応する声調情報を取得し、

ピンイン文字列と声調情報との組合せ及び当該組合せに対応する漢字をそれぞれ対応付けて格納するピンイン変換用辞書を参照し、前記受け付けたピンイン文字列と前記取得した声調情報に基づいて、前記受け付けたピンイン文字列に対応する漢字を取得し、

この取得した漢字と同一の漢字を持つユーザ単語を前記ユーザ単語辞書から検索し、音声情報解析によって取得された声調情報を検索したユーザ単語の声調情報として前記ユーザ単語辞書に登録することを特徴とする中国語文書作成装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、中国語文書の入力に際し、キーボードから入力したピンイン文字列と音声で入力した声調情報を使ってピンイン変換を行なうことにより、中国語の入力操作性の向上を図った中国語文書作成装置及び中国語文書作成方法に関する。

【0002】

【従来の技術】中国語の漢字には、日本語の漢字の読みのような発音表記をアルファベットで表したピンイン文字列が対応している。図1のように、漢字1文字に対応するピンイン文字列は、声母、韻母の対で表現される。中国語におけるピンイン変換は、図2のようにピンイン文字列「g a i」をキーボードから入力し、入力されたピンイン文字列を変換し、対応する変換候補「該、垓、改、蓋、概」を得るものである。しかしながら、日本語のかな漢字変換における同音異義語の問題と同様に、中国語のピンイン変換においても変換候補として複数の同音異義語が出力されるという問題があり、この同音異義語の問題が変換率の低下を招いている。図3に、ピンイン文字列「g a i」及び「q i」に対応する複数の同音異義語の例を示している。

【0003】この問題を解決するために、ピンイン文字列の他に日本語のアクセント情報のような声調情報を入力してピンイン変換を行なうという方法が提案されている。声調情報は、第1声から第5声までの5種類があり、全ての漢字はいずれかの種類の声調情報を持ってい

る。ピンイン文字列の他に声調情報の種類を入力してピンイン変換することにより、図4のように、変換結果として出力する変換候補の数を減らすことができる。

【0004】このような声調情報の種類を入力してピンイン変換を行なう方法としては、特開平8-185396号が提案されていた。しかしながら、この従来技術では、キーボードから、声調情報の種類を示す数字（以下、声調番号という）を入力する（即ち、声調情報の種類である第1声～第5声に対応した1～5の数字を入力する）ものであるため、ピンイン文字列の入力に加えて、声調番号の入力もキーボードから行なう必要があり、ユーザのキー操作が煩雑となるという課題があった。更に、大多数のユーザは、漢字がどのような声調情報の種類を持っているのか正しく認識しておらず、ピンイン文字列と一緒に声調番号もキーボードから入力する方法は、入力速度の低下、正確性の低下を招いているのが現状である。

【0005】

【発明が解決しようとする課題】上述したように、従来技術では、ピンイン変換を行なう際にはピンイン文字列と一緒に声調番号をキーボードから入力する方法が採用されているが、このように声調番号を余分に入力しなければならず、入力操作が煩雑である。また、声調番号を誤って入力してしまうと、希望していない変換候補のみ出力されるため、再度ピンイン文字列と声調番号を入力する必要があるため、声調情報の種類を正しく認識できていないユーザにとっては、反って正確性の低下を招いてしまうという課題があった。

【0006】本発明は、上述した事情を考慮してなされたもので、ピンイン文字列は従来通りキーボードから入力すると共に、声調情報については音声を使って入力することにより入力効率の向上を可能とした中国語文書作成装置及び中国語文書作成方法を提供することを目的とする。

【0007】

【課題を解決するための手段】ピンイン文字列入力手段からピンイン文字列を入力させると共に、音声情報入力手段から音声情報を入力させる。音声情報解析手段により音声情報を解析し、前記ピンイン文字列入力手段から入力されたピンイン文字列に対応する声調情報を取得する。ピンイン変換手段が、ピンイン文字列入力手段から入力されたピンイン文字列と音声情報解析手段により取得された声調情報に基づいてピンイン変換用辞書を参照し対応する漢字を取得する。表示手段が、ピンイン変換手段によって取得された漢字を変換候補として表示する。

【0008】このような中国語文書作成装置及び中国語文書作成方法によれば、音声情報を入力することで声調情報の入力が可能であるため、キー入力はピンイン文字列を入力する場合のみで良いため、声調情報を用いたピ

ンイン変換において、入力操作を容易に且つ迅速に行なうことが可能である。

【0009】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態について説明する。

（第1の実施形態）第1の実施形態は、ピンイン文字列をキーにより入力し、声調情報を音声入力により入力し、この2つの情報からピンイン変換を行なうようにしたものである。

【0010】図5は第1の実施形態に係わる中国語文書作成装置の構成を示したブロック図である。図5において、キーボード入力部1は、キーボードなどで構成されており、ピンイン文字列の入力等、中国語文書作成のための各種入力、編集操作を行うものである。音声入力部2は、マイクなどで構成されており、音声データを入力するためのものである。ポインティングデバイス部3は、ペン入力装置やマウスなどで構成されており、ピンイン変換によって得られた変換候補の選択などを行なうものである。制御部4は、CPUなどで構成されており、キーボード入力部1、音声入力部2、ポインティングデバイス部3から渡される夫々のキーデータ、音声データ、ポインティングデータを使って、後述するピンイン変換部5、音声データ処理部7、出力部10などの各処理部を制御する。ピンイン変換部5は、キーボード入力部1から入力されたピンイン文字列と音声入力部2から入力された音声データにより得られる声調情報から、ピンイン変換用辞書6を参照して漢字に変換し、その変換候補を制御部4に渡す。ピンイン変換用辞書6は、ピンイン文字列に対応する漢字に変換するために必要な情報を収容しているものである。また、ユーザが登録するユーザ単語データもここに記憶される。音声データ処理部7は、音声入力部2から入力された音声データを音声データ処理用辞書8を参照して解析処理することにより、音声データから漢字の声母、韻母、声調情報などの解析結果を取得し、この解析結果を制御部4に渡す。音声データ処理用辞書8は、入力された音声データを解析処理するために必要な情報を収容している。データ記憶部9は、キーボード入力部1から入力されるピンイン文字列、入力されたピンイン文字列に対するピンイン変換に必要なデータや、ピンイン変換を行なった結果得られた変換候補、この変換候補を一覧表示するための候補一覧データ、音声入力部2から入力される音声データ、入力される音声データの解析処理に必要なデータ、ユーザから設定された各種設定情報などを記憶する。出力部10は、CRTやLCDなどからなり、ピンイン変換により得られた変換結果、候補一覧などを表示する。

【0011】次に、第1の実施形態のピンイン変換処理の動作について、図6のフローチャートを参照して説明する。最初に、制御部4は、ユーザからキーボード入力部1を介してピンイン文字列が入力されるのを待ち、ピ

ンイン文字列が入力されるとそのピンイン文字列をデータ記憶部9に記憶する（ステップS101）。引き続き、制御部4は、ピンイン文字列が入力されてから一定時間内に、ユーザから音声入力部2を介して所望する漢字の発音表記が音声データとして入力されるのを待ち、音声データが入力されるとその音声データをデータ記憶部9に記憶する（ステップS102）。制御部4は、ユーザからの音声データの入力がピンイン文字列が入力されてから一定時間内にあったか否かを判断し（ステップS103）、音声の入力があった場合（ステップS103のYES）、入力された音声データを音声データ処理部7へ出力し、音声データから声調情報を取り出す処理を行なわせる（ステップS104）。

【0012】音声データ処理部7は、制御部4から渡された音声データを、音声データ処理用辞書8を参照して解析し、この音声データに対応する漢字の声母、韻母、声調情報などの解析結果を取得する。音声データ処理部7は、得られた解析結果を制御部4に出力する。

【0013】制御部4は、音声データ処理部7から出力された解析結果をチェックし、声調情報が正しく取り出せたか否かを判断する（ステップS105）。すなわち、制御部4は、音声データの解析結果のうち、漢字の声母、韻母の文字列が、ユーザから入力されたピンイン文字列と一致するか否かを判断し、両者が一致する場合に声調情報が正しく取り出せたと判断する。

【0014】この判断により声調情報が正しく取り出せたと判断した場合（ステップS105のYES）、制御部4は、音声データの解析結果から得られる声調情報とデータ記憶部9に記憶されているピンイン文字列とをピンイン変換部5に出力し、ピンイン変換を行なわせる（ステップS106）。ピンイン変換部5は、制御部4から出力されたピンイン文字列と声調情報とからピンイン変換用辞書6を参照することによりピンイン変換を行ない、ピンイン文字列に対応する変換候補を取得する。このときピンイン変換部5は、声調情報を変換候補の候補数の絞込みに用い、声調情報と一致しない漢字は変換候補として出力しないようにする。ピンイン変換部5は、ピンイン変換を行なって得られた変換候補を制御部4に出力する。

【0015】制御部4は、ピンイン変換部5から出力された変換候補をデータ記憶部9に記憶させると共に、当該変換候補を出力部10に出力し、変換候補を表示させる（ステップS108）。ユーザは、出力部10に表示される変換候補の中から所望する変換候補をポインティングデバイス3で指定することにより、変換候補を選択することができる。

【0016】一方、ステップS103の判断においてピンイン文字列が入力されてから一定時間内に音声入力がない場合（ステップS103のNO）、及び、ステップS105の判断において音声データの解析処理の結果、

声調情報が正しく取り出せなかったと判断した場合（ステップS105のNO）、制御部4は、ピンイン文字列のみピンイン変換部5に出力する。ピンイン変換部5は、制御部4から出力されたピンイン文字列のみの情報に基づいて、ピンイン変換用辞書6を参照してピンイン変換を行う（ステップS107）。

【0017】以上説明した動作によりピンイン変換を行なった際の具体例を図7に示している。図7に示すように、ユーザは、ピンイン文字列「gai」と、「gai」の第3声の声調情報をもつ音声を入力することで、その変換候補として「改」を得ることができる。

【0018】このように、ピンイン変換に必要な情報のうち声調情報を音声で入力するようにしたので、ピンイン文字列のみキーボード入力部1から入力すれば良く、入力操作の煩雑化を抑制することができ、入力速度の低下を招くことなく、簡単に声調情報を用いた変換率の高いピンイン変換を行なうことが可能となる。

【0019】なお、上記説明では、単一の漢字をピンイン変換する例を示しているが、複数の漢字からなる単語、或いは、文を一括してピンイン変換することも可能である。即ち、ユーザから、先ず、単語或いは文のピンイン文字列を一括して入力させ、次に、同様の単語或いは文を音声により一括して入力させる。入力された単語或いは文の音声からそれらを構成している各漢字の声調情報を一括して取り出し、この各漢字の声調情報を用いて、先に入力されているピンイン文字列を各漢字毎に変換する。

【0020】また、ユーザからの音声入力を受け付けるタイミングをピンイン文字列入力後の一定時間内としているが、これに限らず、例えば次のピンイン文字列が入力されるまで、音声入力を有効とするとしても良い。

【0021】また、音声で入力する声調情報は、ピンイン文字列として入力される漢字の1文字毎に全て入力する必要はなく、変換候補が多い漢字や、ピンイン変換を行なった際に変換間違いが予想される特定の漢字を入力する場合のみ入力すれば良い。

【0022】また、一般的に中国語の音声認識処理においては、音声データの入力開始時の声母の認識が、その後続く韻母に比べて難しいという問題がある。従って、本実施形態の音声データの解析処理で、声母の認識結果に誤り（声母の認識結果がピンイン文字列と異なる）が比較的多く生じる可能性がある。しかし、声母の認識結果に誤りがあっても、その声母に続く韻母の認識結果が正しい（韻母の認識結果がピンイン文字列と一致する）場合は、これらの声母と韻母を持つ認識結果を正しいものとして判断し、この認識結果により得られた声調情報をピンイン変換に用いるようにしても良い。

【0023】（第2の実施形態）次に、第2の実施形態について説明する。上述した第1の実施形態では、ピンイン変換を行なう際にピンイン文字列と共に音声入力に

て声調情報を入力するようにしたが、これに対して第2の実施形態では、ピンイン変換の結果得られた変換候補に対して音声入力で声調情報を入力することで、その声調情報を有する漢字の変換候補を優先的に表示するようにしたものである。即ち、ピンイン文字列のみを利用してピンイン変換を行なった場合、その変換結果として複数の声調情報を有する漢字が変換候補として得られるため、変換候補の表示ではユーザの所望する声調情報を有する漢字の変換候補が優先的に表示されるとは限らない。このため、変換候補に対して所望する漢字の声調情報を入力できるようにし、入力した声調情報を有する漢字の変換候補を優先的に表示させるようにした。

【0024】第2の実施形態に係わる中国語文書作成装置の構成は、図5に示される第1の実施形態の構成と同じである。第2の実施形態の声調情報の入力による変換候補の表示順位変更処理の動作について、図8を参照して説明する。

【0025】ここで、ピンイン変換部5においてピンイン文字列のみを用いたピンイン変換が行なわれ、出力部10の表示画面上にピンイン変換の変換候補が既に表示された状態にあるものとする。この変換候補の表示は、図9のように、複数の声調情報を有する変換候補のうち表示優先順位の高いもの（ここでは、「該」）が優先的に表示されているものとする。

【0026】先ず、ポインティングデバイス3を介してユーザから出力部10の表示画面上の任意の位置が指定されると、指定された出力部10の表示画面上の指定位置情報を受け付け（ステップS201）、この指定位置情報に基づいて、ユーザから変換候補が指定されたか否かを判断する（ステップS202）。ユーザから変換候補が指定されたと判断した場合（ステップS202のYES）、次に、制御部4は、指定された変換候補についてその他の変換候補があるか否かを判断する（ステップS203）。即ち、制御部4は、指定された変換候補（表示画面上に表示されている変換候補）に代えて表示画面上に表示すべき変換候補がデータ記憶部9に記憶されているか否かを判断する。

【0027】ポインティングデバイス3により指定された変換候補についてその他の変換候補があると判断した場合（ステップS203のYES）、次いで、制御部4は、ポインティングデバイス3による変換候補の指定が、ユーザにより取り消されたか否かを判断する（ステップS204）。

【0028】ユーザによる変換候補の指定が取り消されていない場合（ステップS204のNO）、制御部4は、ユーザから音声入力部2を介して音声データが入力されるのを待ち、音声データが入力されるとそれを受け付け、音声データ処理部7へ出力する（ステップS205）。

【0029】音声データ処理部7は、制御部4から渡さ

れた音声データを、音声データ処理用辞書を参照して解析し、解析結果（漢字の声母、韻母、声調情報など）を取得し、その解析結果を制御部4に出力する（ステップS206）。制御部4は、音声データ処理部7から出力された解析結果をチェックし、声調情報が正しく取り出せたか否かを判断する（ステップS207）。なお、解析結果のチェックの方法は、第1の実施形態で説明した方法と同様である。

【0030】この解析結果のチェックの結果、声調情報が正しく取り出せたと判断した場合（ステップS207のYES）、制御部4は、データ記憶部9に記憶された変換候補の中から、取り出した声調情報を持つものを抽出し、出力部10で現在表示している変換候補に代えて抽出した変換候補を入れ替えて表示するよう制御する（ステップS208）。

【0031】一方、ステップS207の判断において、声調情報が正しく取り出せなかったと判断した場合（ステップS207のNO）、制御部4は、ステップS204に戻り、再び、ポインティングデバイス3による変換候補の指定が、ユーザにより取り消されたか否かを判断する。

【0032】また、ステップS202の判断においてユーザから変換候補が指定されていない判断した場合（ステップS202のNO）、ステップS203の判断においてポインティングデバイス3により指定された変換候補についてその他の変換候補がないと判断した場合（ステップS203のNO）、ステップS204の判断においてユーザによる変換候補の指定が取り消されたと判断した場合（ステップS204のYES）、制御部4は、声調情報の入力による変換候補の表示順位変更処理を終了する。

【0033】以上説明した動作によりピンイン変換を行なった際の具体例を図9、図10に示している。図9は、変換候補「該」に代えて表示すべき変換候補が1つの場合の例を示しており、変換候補「該」を指定して第3声の声調情報を入力することで、「改」が入れ替えて表示される様子を示している。また、図10は、変換候補「該」に代えて表示すべき変換候補が2つの場合の例を示しており、変換候補「該」を指定して第4声の声調情報を入力することで、「蓋」と「概」が一覧で表示される様子を示している。変換候補が複数ある場合はこのように一覧表示し、ユーザに、ポインティングデバイス3を用いて所望の変換候補を選択させるようにしている。

【0034】以上のようにピンイン文字列を用いたピンイン変換を行なって得られた変換候補に対して、音声入力により声調情報を入力し、所望する声調情報を持つ変換候補を優先的に表示させることができる。従って、声調情報を考慮せずに得た変換候補を単純に一覧表示し、その中から所望する変換候補を探す場合に比べ、ユーザにおける変換候補の選択操作を効率良く行なわせること

が可能となる。

【0035】（第3の実施形態）次に、第3の実施形態について説明する。第3の実施形態に係わる中国語文書作成装置の構成は、図5に示される第1の実施形態の構成と同じである。

【0036】第3の実施形態は、ピンイン変換用辞書6に登録されていない単語を新たなユーザ単語として辞書登録するユーザ単語登録に関するものである。ユーザ単語を登録する場合、声調情報が分からない等の理由から、登録するユーザ単語に声調情報を指定しない場合がある。この場合、その後のピンイン変換時にユーザ単語に関する声調情報が入力された際に、この声調情報をユーザ単語の声調情報として自動的に登録するようにしたものである。なお、ここでは、ユーザ単語の登録は、ピンイン変換用辞書6に対して行なわれるものとして説明するが、ピンイン変換用辞書6とは別に、ユーザ単語を登録するためのユーザ辞書を設けても良い。

【0037】次に、第3の実施形態のユーザ単語への声調情報付加処理の動作について、図11のフローチャートを参照して説明する。ここで、変換用辞書6に通常に登録されている単語は、図12のようにピンイン文字列の声母、韻母、及び、声調情報（この例では、声調番号を記憶している）と、これに対応する漢字が対応付けられた状態で記憶されているが、これに対して、図13のようにユーザが登録したユーザ単語には、声調情報が記憶されていないとする。このように、声調情報の記憶されていないユーザ単語に関して、ピンイン変換時に声調情報が入力された際の処理について以下に説明する。

【0038】ステップS301からS307までの処理は、第1の実施形態におけるピンイン変換処理の動作と同様である。即ち、制御部4は、ユーザからキーボード入力部1を介して入力されるピンイン文字列を受け付け、そのピンイン文字列をデータ記憶部9に記憶し（ステップS301）、次いで、ユーザから音声入力部2を介して所望する漢字の発音表記が音声データとして入力されるのを待ち、ピンイン文字列が入力されてから一定時間内に音声データが入力されるとその音声データをデータ記憶部9に記憶する（ステップS302、S303）。音声データ処理部7が、入力された音声データを、音声データ処理用辞書8を参照して解析し、この音声データに対応する漢字の声母、韻母、声調情報などの解析結果を取得し、得られた解析結果を制御部4に出力する（ステップS304）。制御部4は、音声データ処理部7から出力された解析結果をチェックし声調情報が正しく取り出せたか否かを判断する（ステップS305）。制御部4が、声調情報が正しく取り出せたと判断した場合、ピンイン変換部5が、制御部4から出力されたピンイン文字列と声調情報とからピンイン変換用辞書6を参照することによりピンイン変換を行ない、ピンイン文字列に対応する変換候補を取得する（ステップS3

06)。一方、ステップS307の判断においてピンイン文字列が入力されてから一定時間内に音声入力がない場合、及び、ステップS305の判断において音声データの解析処理の結果、声調情報が正しく取り出せなかったと判断した場合のいずれかの場合は、ピンイン変換部5は、ピンイン文字列のみの情報に基づいて、ピンイン変換用辞書6を参照してピンイン変換を行う(ステップS307)。

【0039】次に、制御部4は、ステップS306においてピンイン変換して得られた変換候補と同じ単語で且つ声調情報が記憶されていないユーザ単語が存在するかどうかを判断する(ステップS308)。即ち、制御部4は、ピンイン変換用辞書6を検索し、変換候補のピンイン文字列(声母、韻母)と漢字の組合せと同じ組合せを持つユーザ単語が登録されているかどうかを判断し、変換候補の組合せと同じユーザ単語が登録されている場合、更にそのユーザ単語に声調情報が登録されているかどうかを判断する。

【0040】この結果、変換候補のピンイン文字列と漢字の組合せと同一で、且つ、声調情報の登録されていないユーザ単語が存在すると判断した場合(ステップS308のYES)、制御部4は、ピンイン変換用辞書6に、ユーザ単語の声調情報として、ユーザ単語と同一の組合せの変換候補の声調情報を登録する(ステップS309)。ユーザ単語の声調情報の登録後のピンイン変換用辞書6の例を、図14に示している。

【0041】次いで、制御部4は、ステップS306若しくはステップS307において得られた変換候補を、出力部10にて表示させる(ステップS310)。以上説明した動作の具体例を図15を参照して説明する。

【0042】ピンイン変換用辞書6には、図13のように、ユーザ単語「互聯網」が声調情報の指定のない状態で既に登録されているとする。このとき、ユーザからピンイン文字列「hu lian wang」が入力され、次いで「hu(第4声) lian(第2声) wang(第3声)」と音声データが入力されると、このピンイン文字列と、音声データから得た声調情報から、ピンイン変換を行ない変換候補「互聯網」を得る。ピンイン変換結果後、ピンイン文字列「hu lian wang」と漢字「互聯網」の組合せについてピンイン変換用辞書6を検索し、同一の組合せを持ち且つ声調情報の指定のないユーザ単語を探し出す。図13のようにピンイン変換用辞書6には、上記した組合せに対応し、且つ、声調情報が指定されていないユーザ単語が記憶されているので、このユーザ単語に声調情報を登録する。ユーザ単語に対して声調情報が登録されたピンイン変換用辞書6の様子が図14である。ユーザ単語「互聯網」に対して声調情報「第4声、第2声、第3声」がピンイン変換用辞書6に登録される。

【0043】以上説明した処理によって、ユーザ単語の

登録時に声調情報を登録していなくても、その後のピンイン変換の際に同一の単語に対して音声入力により声調情報を入力すると、入力した声調情報がユーザ単語の登録情報として自動的に登録されるので、ユーザによるユーザ単語登録操作が簡単となる。

【0044】なお、上述した各実施形態において記載した手法は、コンピュータに実行させることのできるプログラムとして、例えば磁気ディスク(ハードディスク等)、光ディスク(CD-ROM、DVD等)、半導体メモリ等の記憶媒体に書きこんで各種装置に適用したり、通信媒体により伝送して各種装置に適用することも可能である。本装置を実現するコンピュータは、記憶媒体に記憶されたプログラムを読み込み、このプログラムによって動作が制御されることにより、上述した処理を実行する。

【0045】

【発明の効果】以上詳述したように本発明によれば、声調情報を用いてピンイン変換を行なうに際し、声調情報の入力を所望する漢字の音声情報を入力することで行なうことが可能であるため、ピンイン文字列のみキー入力を行えば良く、ユーザによる入力操作を容易且つ迅速に行なうことが可能である。

【0046】また、ピンイン文字列のみでピンイン変換を行なった後、所定の変換候補に対して、音声情報を入力することにより声調情報を入力し、所望する漢字の変換候補を優先的に表示することが可能であるため、ユーザによる変換候補の選択操作を容易且つ迅速に行なうことが可能である。

【0047】また、ピンイン変換の際に音声情報により入力された声調情報を、ユーザ単語辞書に登録されたユーザ単語の声調情報として登録するようにしたので、ユーザは声調情報が分からない場合はユーザ単語辞書に登録しなくても良くなり、ユーザ単語登録の際の利便性が向上する。

【図面の簡単な説明】

【図1】漢字とピンイン文字列との対応例を示した図。

【図2】ピンイン文字列のみでピンイン変換を行なう際の動作を示した図。

【図3】同じピンイン文字列を持つ漢字例示した図。

【図4】同じピンイン文字列を持つ漢字の声調情報を示した図。

【図5】本発明に係わる中国語文書作成装置の構成を示した図。

【図6】第1の実施形態のピンイン変換処理の動作を示すフローチャート。

【図7】第1の実施形態のピンイン変換処理の具体例を示した図。

【図8】第2の実施形態の声調情報の入力による変換候補の表示順位変換処理の動作を示すフローチャート。

【図9】第2の実施形態の声調情報の入力による変換候



補の表示順位変換処理の第1の具体例を示した図。

【図10】第2の実施形態の声調情報の入力による変換候補の表示順位変換処理の第2の具体例を示した図。

【図11】本発明の第3の実施形態のユーザ単語への声調情報付加処理の動作を示すフローチャート。

【図12】変換用辞書6に格納された通常の単語の格納例を示した図。

【図13】変換用辞書6に格納された声調情報の付加されていないユーザ単語の格納例を示した図。

【図14】変換用辞書6に格納された声調情報の付加されたユーザ単語の格納例を示した図。

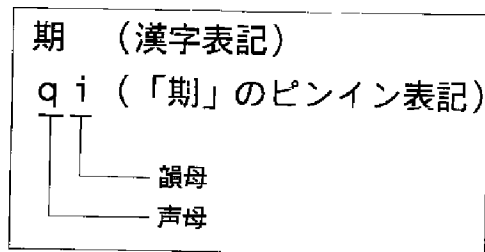
【図15】本発明の第3の実施形態のユーザ単語への声

調情報付加処理の具体例を示した図。

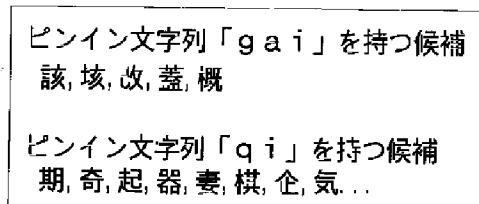
【符号の説明】

- 1…キーボード入力部
- 2…音声入力部
- 3…ポインティングデバイス
- 4…制御部
- 5…ピンイン変換部
- 6…ピンイン変換用辞書
- 7…音声データ処理部
- 8…音声データ処理用辞書
- 9…データ記憶部
- 10…出力部

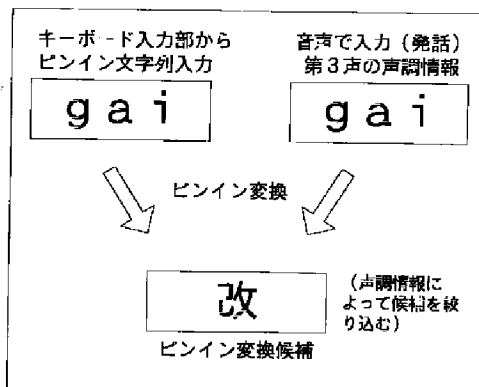
【図1】



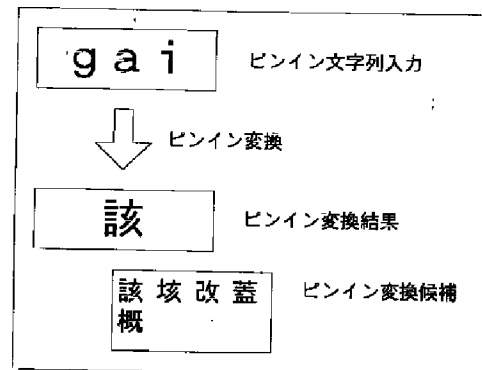
【図3】



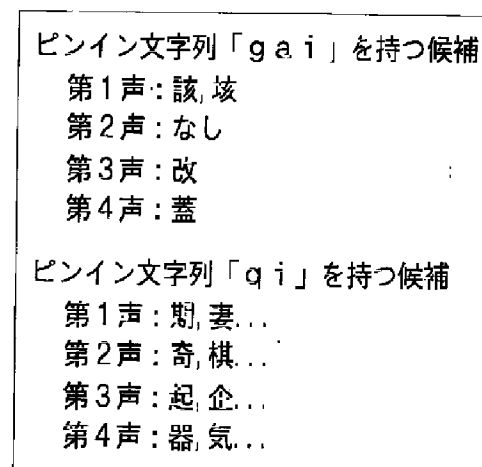
【図7】



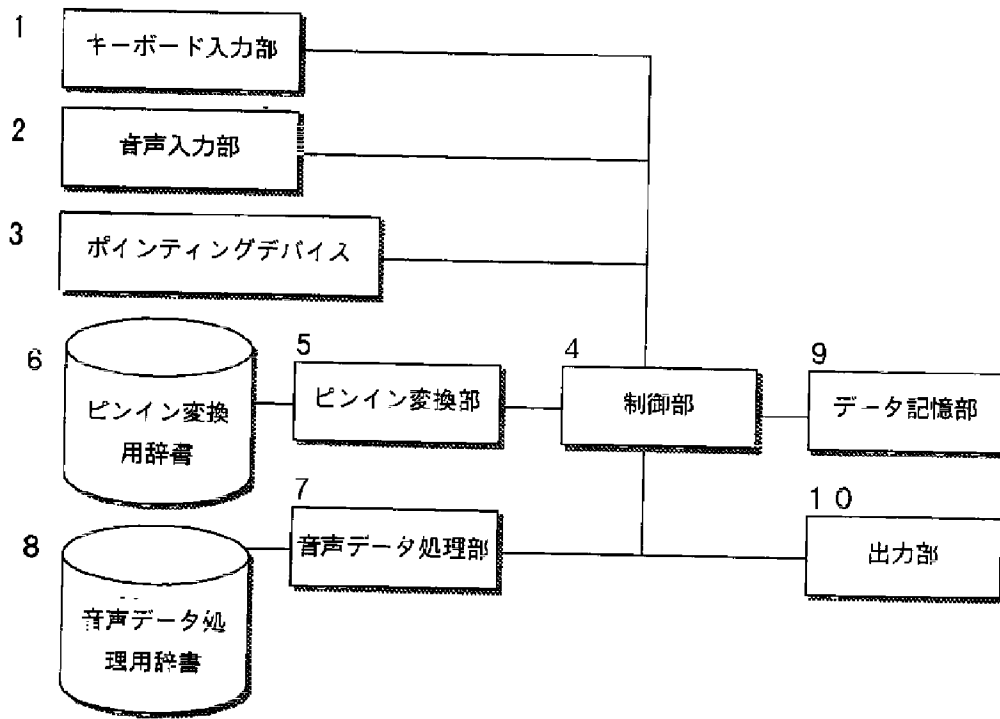
【図2】



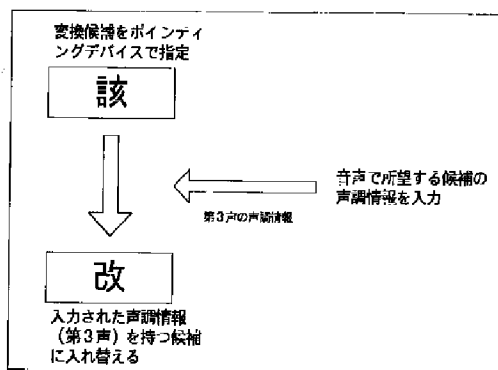
【図4】



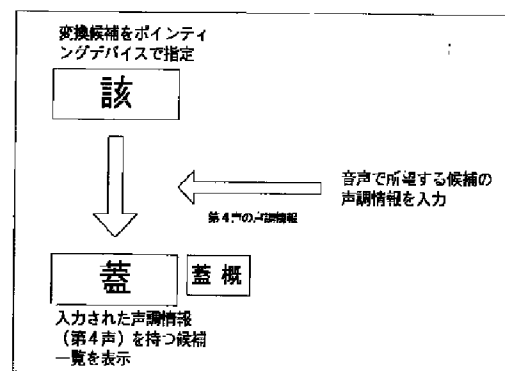
【図5】



【図9】



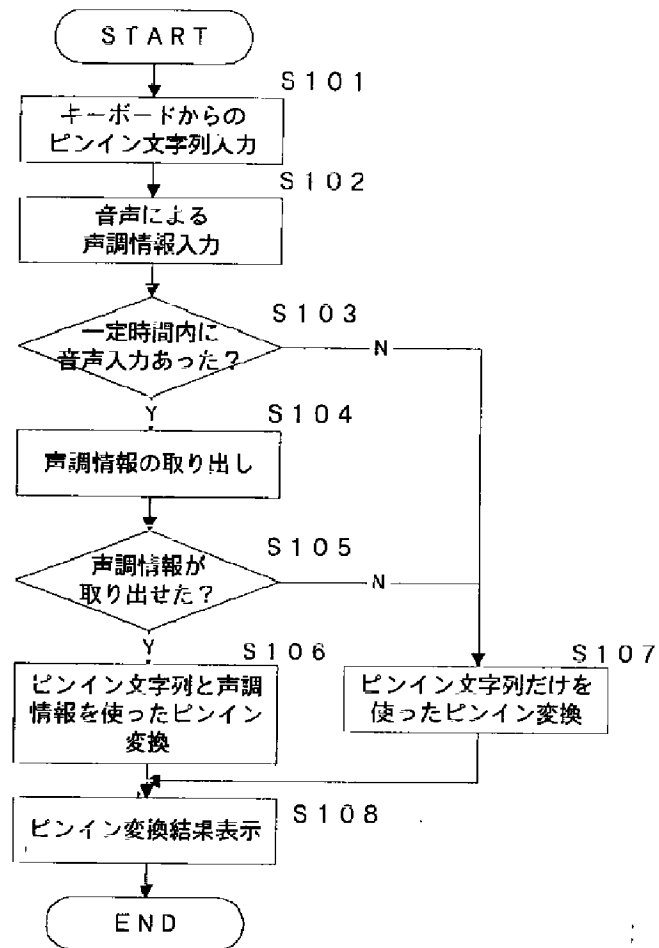
【図10】



【図13】

声母	韻母	声調番号	漢字	声母	韻母	声調番号	漢字	声母	韻母	声調番号	漢字
h	u		互	l	ian		聯	w	ang		網

【図6】



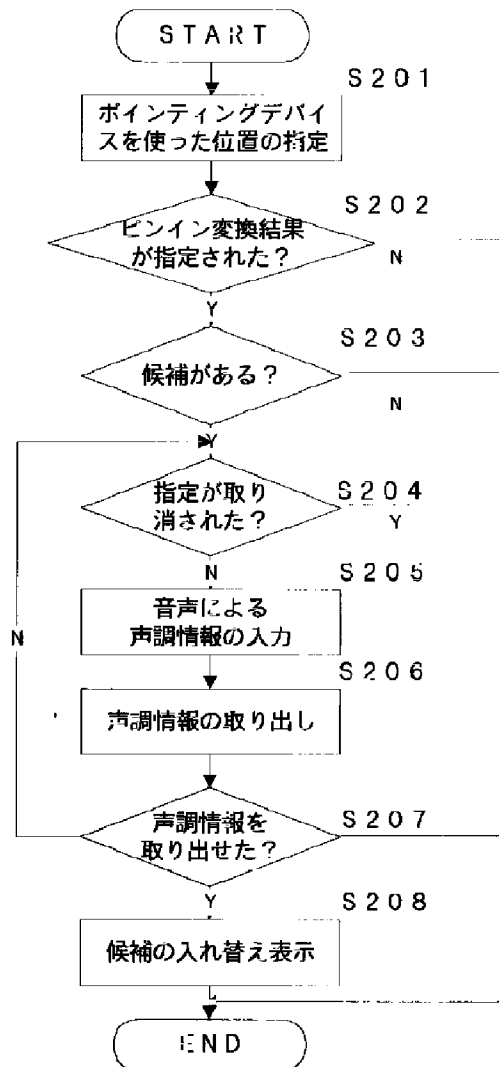
【図12】

声母	韻母	声調番号	漢字
g	ai	1	該
...			
q	i	1	期
q	i	1	妻
...			
q	i	4	器
...			

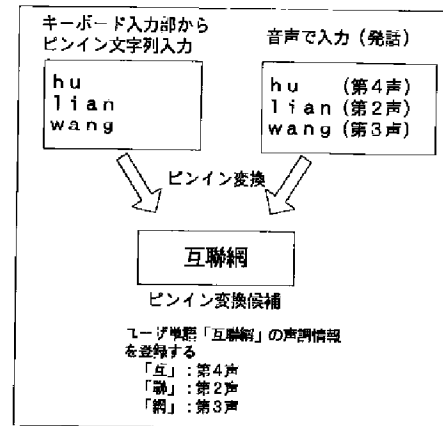
【図14】

声母	韻母	声調番号	漢字	声母	韻母	声調番号	漢字	声母	韻母	声調番号	漢字
h	u	4	互	1	ian	2	聯	w	ang	3	翁

【図8】



【図15】



【図11】

